WUCSE-2005-60

Functional Optimization Models for Active Queue Management

Authors: Chen, Yixin

December 26, 2005

Corresponding Author: chen@cse.wustl.edu

Functional Optimization Models for Active Queue Management

Yixin Chen

Department of Computer Science and Engineering
Washington University in St Louis
1 Brookings Drive
St Louis, MO 63130, USA
`chen@cse.wustl.edu`
URL: `http://www.cse.wustl.edu/~chen/`

Abstract. Active Queue Management (AQM) is an important problem in networking. In this paper, we propose a general functional optimization model for designing AQM schemes. Unlike the previous static function optimization models based on the artificial notion of utility function, the proposed *dynamic functional* optimization formulation allows us to directly characterize the desirable system behavior of AQM and design AQM schemes to optimally control the dynamic behavior of the system. Such a formulation also allows adaptive control which enables the AQM scheme to continuously adapt to dynamic changes of networking conditions. In this paper, we present the Pontryagin minimum principle, a necessary condition, for the functional optimization model of AQM with TCP AIMD congestion control. As an example, we investigate a queuing stability criteria and apply the necessary conditions to optimize the functional model.

1 Introduction

Active Queue Management (AQM) is an important problem in networking. This paper presents a functional optimization model for designing AQM schemes.

Starting from the initial work of F. Kelly et al. [6, 8, 7], there have been great interests in a utility *function optimization model* for designing AQM schemes. The model assumes that each user has a *utility function* which the user tries to maximize. The objective of the system is to maximize the summation of user utility functions subject to the resource (link capacity) constraints. This constrained function optimization problems is then solved using either penalty function methods [6, 8, 7, 9, 11] or duality model methods [13, 12, 2].

Various AQM schemes [10, 2], are derived so that each end hosts and each router can carry out a distributed algorithms in such a way that the whole system asymptotically solves the function optimization problem. Assuming that the utility functions are continuously differentiable, strictly concave, and monotonically increasing, various stability results have been established which show that the system will asymptotically converge to the optimal solution as time approaches infinity.

There are two major problems with such a function optimization model. First, the utility function may not be a realistic objective. The utility function is artificially introduced as the goal of the AQM schemes, and does not necessarily characterize the desirable objective of networking systems. Second, the function optimization model is static (assuming no flow will die and no new flow will join) and the stability results only analyze the asymptotic behavior of the system. However, in real applications, networking conditions is constantly changing, the static assumption is violated, and the stability results may not be useful.

To address the limitations of existing methods, we propose in this paper a new dynamic functional optimization model of AQM. A *functional* is a mapping from the curves into a numerical value. We take end host source rates and queuing lengths as *state variables* and dropping/marking probabilities as *control variables*. All these system variables evolve with time and correspond to bundles of curves that characterize the system behavior. The system objective of AQM are defined as a *functional* of these system curves. In addition, the relationship between state variables and control variables are specified in *system control functions*. Based on the theory of calculus of variations, we develop the Pontryagin minimum principle for the proposed functional optimization model of AQM with TCP AIMD (Additive-Increase, Multiplicative-Decrease) congestion control mechanism.

The functional optimization model proposed in this paper is significant for two reasons. First, it provides us the flexibility and accuracy to directly specify the design objective of AQM schemes as functional of dynamic system variables. We can also incorporate addition side constraints to further define the desirable operating regions of the system. Second, the functional optimization model enables the AQM scheme to continuously adapt to changes of networking conditions through adaptive control, so that the system always follows the optimal control trajectory based on current networking conditions.

The paper is organized as follows. In Section 2, we overview related previous work in AQM. In Section 3, we present a general framework of functional optimization models for AQM. Section 4 presents Pontryagin minimum principle, a necessary condition for optimal solutions to the AQM model with TCP AIMD congestion control. In Section 5, we present an example of applying the minimum principle to a queuing stability criteria. Section 6 concludes the paper.

2 Previous Work

Previous work in AQM, including RED, FRED, SRED, BLUE, stochastic blue, etc, adapts the dropping/marking probability using heuristic rules in order to achieve various system objectives, including fairness, high link utilization, stabilized queue length, etc. The newly developed PAQM [5] scheme predicts future traffic using a LMMSE predictor and adapts the dropping probability using a controller stabilizing the queue length based on the traffic prediction.

We briefly review previous work on the utility function optimization model in the following. Adopting the following system model described by F. Kelly and

R. Srikant in [6, 11], consider a network with a set $\mathcal{L}$ of links and let C_l be the capacity of link l, for $l = 1, \cdots, |\mathcal{L}|$. Let a route r be a non-empty subset of $\mathcal{L}$, and let $\mathcal{R}$ be the set of possible routes. Let S be a $|\mathcal{R}| \times |\mathcal{L}|$ matrix, set $S_{rl} = 1$ if $l \in r$; so that route r traverses link l and set $S_{rl} = 0$ otherwise. Let route r generates at rate x_r.

The rate x_r is assumed to have a utility $U_r(x_r)$ to user r. Assume that $U_r(\cdot)$ is a continuously differentiable, strictly concave, increasing function in the interval $(0, \infty)$ and assume that $U_r(x_r)$ is unbounded as x_r approaches 0 to ensure that each user gets some non-zero throughput. Examples of such a function include $log(x_r)$ and $-1/x_r$. Let $C = (C_l, l \in \mathcal{L})$, the optimal rates for this network can now be obtained by solving the following constrained *function optimization* problem:

$$\begin{aligned} \text{maximize}_x \ & \sum_{r=1}^{|\mathcal{R}|} U_r(x_r) \\ \text{subject to} \ & S^T x \leq C \end{aligned} \tag{1}$$

There are two major approaches to solve this constrained optimization problem: penalty function methods [6, 8, 7, 9, 11] or duality model methods [13, 12, 2]. Various AQM schemes, such as AVQ [10] and REM [2], are proposed based on this model and in essence carry out a distributed optimization algorithm solving (1). AVQ maintains a virtual queue with adjustable size and uses input rate as congestion index. Asymptotic stability of AVQ is proved based on a single link model. REM uses a quantity called price to feedback the congestion index to end users. Local asymptotic stability of REM is proved based on a multiple link model.

The system goal of this model, i.e. utility function, is artificially introduced in order to establish the model. Theoretic analysis typically first derive the distributed algorithms for end hosts and routers and try to fit the distributed algorithm into end hosts' congestion control scheme. Since most end hosts are using the AIMD algorithm of TCP, the utility function has to take a particular form in order for TCP to be the local algorithm for end users. Therefore, the utility function objective may not be realistic and there exists a major discrepancy between actual AQM system objective and the utility function objective. The utility function model may not necessarily characterize the desirable system behavior.

Various asymptotic stability results of the distributed algorithms are proved [6, 8, 7, 9, 11, 2, 15]. The asymptotic stability results ensures that the system will approach the optimal solution to the utility function optimization model as time approaches infinity. The assumption of the stability property is that the system is static, i.e. no route will die and no new route will join, which is not true in real computer networks.

3 Functional Optimization Model of AQM

From the point view of dynamic optimization, we proposed the following *functional optimization* formulation for AQM. The dynamic system Υ evolves with time t and contains both *state variables*, *control variables*, and *control functions* governing the dynamics of the system.

State of system Υ can be described by the input rate of each user and the queue length at each router. Therefore, system Ψ has the following state variables at time t: $x_r(t)$ for $r = 1, \cdots, |\mathcal{R}|$, and $\omega_l(t)$ for $l = 1, \cdots, |\mathcal{L}|$, where $\omega_l(t)$ denotes the queue length of router l at time t. More congestion indexes can also be added into the state variables. Since we assume the congestion control policy at end hosts is fixed, the only control variables of the system is $p_l(t)$, the dropping probability of router l at time t, for $l = 1, \cdots, |\mathcal{L}|$. If we also want to design the congestion control policy at end hosts, its control parameters should also be included in the control variables and be determined by optimization.

Therefore, we have the following *state vectors* at time t:

$$x(t) = (x_1(t), x_2(t), \cdots, x_{|\mathcal{R}|}(t))^T, \tag{2}$$

$$\text{and} \quad \omega(t) = (\omega_1(t), \omega_2(t), \cdots, \omega_{|\mathcal{L}|}(t))^T \tag{3}$$

and the following *control vector*:

$$p(t) = (p_1(t), p_2(t), \cdots, p_{|\mathcal{L}|}(t))^T. \tag{4}$$

In general, we assume that system Υ is governed by *control functions*:

$$\Delta x(t) = G_x(x(t), \omega(t), p(t), t), \tag{5}$$

$$\text{and} \quad \Delta\omega(t) = G_\omega(x(t), \omega(t), p(t), t), \tag{6}$$

where the general operator Δ means differentiation for continuous-time systems and difference for discrete-time systems.

Since each state variable or control variable corresponds to a curve (continuous or discrete) along the time dimension, the system is composed of several bundles (a bundle is a vector of curves). Let x denote the bundle of $x(t)$, ω denote the bundle of $\omega(t)$, and p denote the bundle of $p(t)$, the objective of AQM is to find the bundle p which forces the system, from any given initial state $x(0)$ and $\omega(0)$, to evolve while minimizing the following functional:

$$\text{minimize}_p \qquad J[x, \omega, p], \tag{7}$$

subject to the system control dynamics in (5) and (6), where J is a functional (a mapping from a set of curves to a numerical value) defining the performance objective of an AQM scheme. Typical goals of AQM schemes include stabilizing queue lengths, maximizing link utilization, minimizing packet loss ratios, and providing fairness guarantees. Such criteria can be readily defined in the functional J.

Unlike the formulation of utility function optimization, which is unrealistic and only concerns about the asymptotic behavior, the functional formulation provides the power of dynamically specifying the system objective on real-time behavior, without recourse to artificial utility functions. In addition, based on the optimal control function obtained from the functional optimization model, the routers can perform continuous replanning when a existing route terminates or a new route joins. Such an adaptive control process enables the AQM scheme to continuously adapt to changes of networking conditions and follow the optimal control trajectory based on current networking conditions.

Under this general functional analysis formulation, there are two major design issues to be considered.

a) Discrete-time system vs. continuous-time system. We study discrete-time system in this paper, because in real implementation, a router cannot continuously adapt the control variables (dropping probabilities). Instead, the control variables are updated at discrete time points. Moreover, optimal continuous-time control can be obtained from discrete-time results by limiting the time unit to infinitesimal.
b) Dynamics of end host traffic rate. The AIMD (Additive-Increase, Multiplicative-Decrease) scheme of TCP is the most widely used congestion control policy in the Internet. Other schemes, such as MIMD, are also actively studied. PAQM [5] provides a novel scheme in which the traffic rate is not modeled by end-host protocol, but instead predicted based on the long-range dependence property of Internet traffic. We study the AIMD dynamics of TCP in this paper.

4 Pontryagin Minimum Principle for AQM with TCP

Calculus of variations is the major tool for attacking functional optimization problems. In this section, we derive the Pontryagin minimum principle for discrete system $\varUpsilon$ with TCP AIMD schemes at end hosts. Following the classic formulation of discrete calculus of variations [3], for a system with planning horizon N, we need to find the discrete-time bundle p of dropping probabilities in order to minimize the functional:

$$J[x,\omega,p] = \sum_{t=0}^{N+1} F(t, x(t), \omega(t), p(t)) \tag{8}$$

subject to system control functions, where F is a scalar function which has continuous derivatives with respect to elements in $x(t), \omega(t), p(t)$.

It has been derived from numerous papers [12, 14, 9] that the AIMD algorithm of TCP can be modeled as,

$$\varDelta x_r(t) = x_r(t+1) - x_r(t) = \frac{1}{d^2} - \beta x_r^2(t) \sum_{l=1}^{|\mathcal{L}|} S_{rl} p_l(t), \quad \forall r = 1, \cdots, |\mathcal{R}| \tag{9}$$

based on the simplifying assumption that all routes have round-trip delay d and that link losses are independent [9]. Constant β is derived as 1/2 in [12], ln2 in [9], and 2/3 in [14]. For simplicity, we set $\beta = 1/2$ in this paper.

The dynamics of the queue length at each router are as follows:

$$\Delta\omega_l(t) = \omega_l(t+1) - \omega_l(t) = (1 - p_l(t)) \sum_{r=1}^{|\mathcal{R}|} S_{rl} x_r(t) - C_l, \quad \forall l = 1, \cdots, |\mathcal{L}| \tag{10}$$

Summing up, the functional optimization model for AQM with AIMD congestion control of TCP can formulated as follows:
$SYSTEM\ \Upsilon(x, \omega, p)$:

$$\text{minimize}_p \qquad J[x, \omega, p] = \sum_{t=0}^{N+1} F(t, x(t), \omega(t), p(t)), \tag{11}$$

subject to:

$$\begin{aligned} x_r(t+1) &= \Psi_r(x(t), p(t)) \\ &= x_r(t) + \frac{1}{d^2} - \frac{x_r^2(t)}{2} \sum_{l=1}^{|\mathcal{L}|} S_{rl} p_l(t), \quad \forall r = 1, \cdots, |\mathcal{R}| \end{aligned} \tag{12}$$

$$\begin{aligned} \omega_l(t+1) &= \Omega_l(\omega(t), x(t), p(t)) \\ &= \omega_l(t) + (1 - p_l(t)) \sum_{r=1}^{|\mathcal{R}|} S_{rl} x_r(t) - C_l, \quad \forall l = 1, \cdots, |\mathcal{L}| \end{aligned} \tag{13}$$

where the initial states $x(0)$ and $\omega(0)$ are known.

We apply the method of calculus of variations to solve this problem. We introduce the Lagrange multiplier to formulate the Hamiltonian function of $SYSTEM\ \Upsilon(x, \omega, p)$, and derive the Pontryagin minimum principle, a necessary condition for optimal solutions to the system.

We introduce the following Lagrangian function:

$$\begin{aligned} \Gamma = \sum_{t=0}^{N+1} \Big(F(t, x(t), \omega(t), p(t)) \Big) + \sum_{t=0}^{N} \Big(& \lambda^T(t) [\Psi(x(t), p(t)) - x(t+1)] \\ & + \mu^T(t) [\Omega(\omega(t), x(t), p(t)) - \omega(t+1)] \Big) \end{aligned} \tag{14}$$

where $\lambda(t)$ is an $|\mathcal{R}|$-element vector and $\mu(t)$ is an $|\mathcal{L}|$-element vector. Both $\lambda(t)$ and $\mu(t)$ serve as Lagrange multipliers in (14). Ψ and Ω are vectors of functions defined in (12) and (13), respectively.

We proceed to define the Hamiltonian function of the system based on the Lagrangian formulation.

Definition 1. *The Hamiltonian function of $SYSTEM\ \Upsilon(x, \omega, p)$ is defined as:*

$$\begin{aligned} \mathcal{H}(t, x(t), \omega(t), p(t)) = F(t, x(t), \omega(t), p(t)) + \lambda^T(t) \Psi(x(t), p(t)) \\ + \mu^T(t) \Omega(\omega(t), x(t), p(t)), \quad t = 0, 1, \cdots, N \end{aligned}$$

Based on variational calculus theory, we present in the following theorem a necessary condition for optimal solutions to $SYSTEM\ \Upsilon(x, \omega, p)$. Define $\odot$ to be the vector multiplication operator (for two vectors u and v with same length k, $u \odot v = z$, where $z_i = u_i * v_i$ for $i = 1, 2, \cdots, k$), the necessary condition is stated as follows:

Theorem 1. *Pontryagin minimum principle for $SYSTEM\ \Upsilon(x, \omega, p)$. For any solution to $SYSTEM\ \Upsilon(x, \omega, p)$, there must exist bundles λ and μ so that the following equations are satisfied:*

$$\lambda(t-1) = \lambda(t) - \lambda(t) \odot x(t) \odot [Sp(t)] + S[\mu(t) - \mu(t) \odot p(t)] + \nabla_{x(t)} F, \quad t = 1, 2, \cdots, N \tag{15}$$

$$\mu(t-1) = \mu(t) + \nabla_{\omega(t)} F, \quad t = 1, 2, \cdots, N \tag{16}$$

$$0 = -\frac{1}{2} S^T [\lambda(t) \odot x(t) \odot x(t)] - \mu(t) \odot [S^T x(t)] + \nabla_{p(t)} F, \quad t = 0, 1, \cdots, N \tag{17}$$

$$\lambda(N) = \nabla_{x(N+1)} F \quad \text{and} \quad \mu(N) = \nabla_{\omega(N+1)} F \tag{18}$$

Proof. According to the theory of Lagrange multipliers in continuous space [1], the necessary conditions for solving $SYSTEM(x, \omega, p)$ is that x, ω, p should extremize Γ. We extremize Γ with respect to $x(t)$:

$$\nabla_{x(t)} \Gamma = -\lambda(t-1) + \nabla_{x(t)} \mathcal{H}(t, x(t), \omega(t), p(t)) \tag{19}$$

$$= -\lambda(t-1) + \nabla_{x(t)} \Big(F(t, x(t), \omega(t), p(t)) + \lambda^T(t) \Psi(x(t), p(t)) + \mu^T(t) \Omega(\omega(t), x(t), p(t)) \Big) = 0, \quad t = 1, 2, \cdots, N \tag{20}$$

$$\nabla_{x(N+1)} \Gamma = -\lambda(N) + \nabla_{x(N+1)} F = 0 \tag{21}$$

with respect to $\omega(t)$:

$$\nabla_{\omega(t)} \Gamma = -\mu(t-1) + \nabla_{\omega(t)} \mathcal{H}(t, x(t), \omega(t), p(t)) \tag{22}$$

$$= -\mu(t-1) + \nabla_{\omega(t)} \Big(F(t, x(t), \omega(t), p(t)) + \mu^T(t) \Omega(\omega(t), x(t), p(t)) \Big) = 0, \quad t = 1, 2, \cdots, N \tag{23}$$

$$\nabla_{\omega(N+1)} \Gamma = -\mu(N) + \nabla_{\omega(N+1)} F = 0 \tag{24}$$

and with respect to $p(t)$:

$$\nabla_{p(t)} \Gamma = \nabla_{p(t)} \mathcal{H}(t, x(t), \omega(t), p(t)) \tag{25}$$

$$= \nabla_{p(t)} \Big(F(t, x(t), \omega(t), p(t)) + \lambda^T(t) \Psi(x(t), p(t)) + \mu^T(t) \Omega(\omega(t), x(t), p(t)) \Big) = 0, \quad t = 0, 1, \cdots, N \tag{26}$$

From (12) and (13), for $x(t)$, we have:

$$\bigtriangledown_{x(t)}[\lambda^T(t)\Psi(x(t),p(t))] = \lambda(t) - \lambda(t)\odot x(t)\odot[Sp(t)] \quad (27)$$

$$\bigtriangledown_{x(t)}[\mu^T(t)\Omega(\omega(t),x(t),p(t))] = S[\mu(t) - \mu(t)\odot p(t))] \quad (28)$$

For $\omega(t)$, we have:

$$\bigtriangledown_{\omega(t)}[\mu^T(t)\Omega(\omega(t),x(t),p(t))] = \mu(t), \quad (29)$$

and, for $p(t)$, we have:

$$\bigtriangledown_{p(t)}[\lambda^T(t)\Psi(x(t),p(t))] = -\frac{1}{2}S^T[\lambda(t)\odot x(t)\odot x(t)] \quad (30)$$

$$\bigtriangledown_{p(t)}[\mu^T(t)\Omega(\omega(t),x(t),p(t))] = -\mu(t)\odot[S^Tx(t)] \quad (31)$$

Substituting the derivatives in (27) - (31) into (20), (23), and (26), we get:

$$\lambda(t-1) = \lambda(t) - \lambda(t)\odot x(t)\odot[Sp(t)] + S[\mu(t) - \mu(t)\odot p(t)) + \bigtriangledown_{x(t)}F \quad (32)$$

$$\mu(t-1) = \mu(t) + \bigtriangledown_{\omega(t)}F \quad (33)$$

$$0 = -\frac{1}{2}S^T[\lambda(t)\odot x(t)\odot x(t)] - \mu(t)\odot[S^Tx(t)] + \bigtriangledown_{p(t)}F \quad (34)$$

The theorem is proved after combining (21), (24), (32), (33), and (34). ■

Theorem 1 is important in that, combined with the system control functions (12) and (13), it provides a system of first-order difference equations whose solution will in general lead to the optimal solution of $SYSTEM\ \Upsilon(x,\omega,p)$. We present an example of applying Theorem 1 in the next section.

5 An Example: Queuing Stability Criteria

Stability of queuing is a very important objective of AQM. A stable queue length enables full utilization of link capacity while not incurring excessive packet loss. To stabilize queue lengths, we define the following functional objective for $SYSTEM\ \Upsilon(x,\omega,p)$:

$$J[x,\omega,p] = \sum_{t=0}^{N+1}(\omega(t)-\omega^*)^T(\omega(t)-\omega^*) \quad (35)$$

where ω^* is a constant denoting the desirable queue length.

We apply Theorem 1 to this criteria. Let $[x]_b^a = x$ if $a \le x \le b$, $[x]_b^a = a$ if $x > a$, and $[x]_b^a = b$ if $x < b$, we have the following result:

Theorem 2. *For $SYSTEM\ \Upsilon(x,\omega,p)$ with functional in (35), the optimal control of dropping probability $p(t)$ is, $\forall t = 0,1,\cdots,N$:*

$$p_l(t) = \left[1 - \frac{\omega^* + C_l - \omega_l(t)}{\sum_{r=1}^{|\mathcal{R}|}S_{rl}x_r(t)}\right]_0^1, \quad l = 1,2,\cdots,|\mathcal{L}| \quad (36)$$

Proof. Since in this example,

$$F(t, x(t), \omega(t), p(t)) = (\omega(t) - \omega^*)^T(\omega(t) - \omega^*) \tag{37}$$

we have:

$$\bigtriangledown_{x(t)}(F(t, x(t), \omega(t), p(t)) = 0 \tag{38}$$
$$\bigtriangledown_{\omega(t)}(F(t, x(t), \omega(t), p(t)) = 2(\omega(t) - \omega^*) \tag{39}$$
$$\bigtriangledown_{p(t)}(F(t, x(t), \omega(t), p(t)) = 0 \tag{40}$$

Substituting (38), (39) and (40) into the necessary conditions of Theorem 1 yields:

$$\lambda(t-1) = \lambda(t) - \lambda(t) \odot x(t) \odot [Sp(t)] + S[\mu(t) - \mu(t) \odot p(t)], \quad t = 1, 2, \cdots, N \tag{41}$$
$$\mu(t-1) = \mu(t) + 2(\omega(t) - \omega^*), \quad t = 1, 2, \cdots, N \tag{42}$$
$$0 = -\frac{1}{2}S^T[\lambda(t) \odot x(t) \odot x(t)] - \mu(t) \odot [S^T x(t)], \quad t = 0, 1, \cdots, N \tag{43}$$
$$\lambda(N) = 0, \quad \text{and} \quad \mu(N) = 2(\omega(N+1) - \omega^*) \tag{44}$$

In order to solve the control variable $p(t)$, we substitute the control function (13) into (42) and expand (42) in scalar form as:

$$\mu_l(t) = \mu_l(t+1) + 2(\omega_l(t) + (1 - p_l(t))\sum_{r=1}^{|\mathcal{R}|} S_{rl}x_r(t) - C_l - \omega^*) \quad l = 1, 2, \cdots, |\mathcal{L}|, t = 0, 1, \cdots, N-1 \tag{45}$$

which can be rewritten as, $\forall l = 1, 2, \cdots, |\mathcal{L}|, t = 0, 1, \cdots, N-1$:

$$p_l(t) = 1 - \frac{\frac{\mu_l(t) - \mu_l(t+1)}{2} + \omega^* + C_l - \omega_l(t)}{\sum_{r=1}^{|\mathcal{R}|} S_{rl}x_r(t)} \tag{46}$$

Similarly, substituting (13) into (44) gets, $\forall l = 1, 2, \cdots, |\mathcal{L}|$:

$$p_l(N) = 1 - \frac{\frac{\mu_l(N)}{2} + \omega^* + C_l - \omega_l(N)}{\sum_{r=1}^{|\mathcal{R}|} S_{rl}x_r(N)} \tag{47}$$

The optimal control of dropping probabilities $p(t)$ can be solved using a backward iterative process. At the first iteration, we start from $\lambda(N) = 0$, substitute it into (43) to yield $\mu(N) = 0$, which is then substituted into (47) to get:

$$p_l(N) = \left[1 - \frac{\omega^* + C_l - \omega_l(N)}{\sum_{r=1}^{|\mathcal{R}|} S_{rl}x_r(N)}\right]_0^1, \quad l = 1, 2, \cdots, |\mathcal{L}| \tag{48}$$

The dropping probability is bounded in $[0, 1]$ and is set to be on the boundary when out of feasible range [4].

We then substitute $\lambda(N) = 0$ and $\mu(N) = 0$ into (41) and get $\lambda(N-1) = 0$ to start the second iteration. At the second iteration, we substitute $\lambda(N-1) = 0$ into (43) to yield $\mu(N-1) = 0$, substitute $\mu(N-1) = 0$ into (46) to get $p(N-1)$, substitute $\lambda(N-1) = 0$ and $\mu(N-1) = 0$ into (41) and get $\lambda(N-2) = 0$ to start the next iteration. Repeat the backward iteration until $t = 0$ will yield the optimal control functions for $p(t), t = 0, 1, \cdots, N$ in (36). ■

To implement (36), at time t, a router $l, l = 1, 2, \cdots, |\mathcal{L}|$ can observe C_l and $\omega_l(t)$. Although router l does not know all the $x_r(t)$ for $r = 1, 2, \cdots, |\mathcal{R}|$, but all it needs to know is the aggregated traffic $\sum_{r=1}^{|\mathcal{R}|} S_{rl} x_r(t)$, which can be observed at the router. The queuing stability criteria studied is an illustration of applying Theorem 1. We may need more sophisticated functional criteria which can take more effective or multiple criteria, e.g. low loss ratio, fairness, etc., into consideration.

6 Conclusions

In this paper, we have proposed a general functional optimization model for designing AQM schemes. The proposed *dynamic functional* optimization formulation models AQM as an optimal control problem and enables dynamic optimization of system behavior based system states, including input rates and queue length. Such a formulation also allows dynamic replanning which enables the AQM scheme to continuously adapt to changes of networking conditions.

We have further developed the Pontryagin minimum principle, a necessary condition, for the functional optimization model of AQM with TCP AIMD congestion control. As an example, we have presented a queuing stability criteria and apply the necessary conditions to optimize the functional model.

The proposed functional optimization model is general and flexible. There are lots of interesting open issues to explored. We need to design more sophisticated functional formulation to enforce the system to be controlled in the most desirable way. We can also study multi-objective formulations through which multiple conflicting system objectives can be balanced. Another important extension is incorporation of additional side constraints to the functional optimization model. Variational theory can handle Lagrange constraints and isoperimetric constraints [3]. We can specify constraints on various measures, such as link utilization, packet loss ratio, and queuing delay. Side constraints will allow us to define more precisely the desirable operating regions of the system, and rule out unacceptable behavior of the system.

References

1. T. M. Apostol. *Mathematical Analysis*. Addison-Wesley, Boston, MA, 1957.
2. A. Athuraliya, *et al.* Rem: Active queue management. *IEEE Network Magazine*, 15(3):48–53, 2001.

3. J. A. Cadzow. Discrete calculus of variations. *Int. J. Control*, 11:393–407, 1970.
4. S. E. Dreyfus. *Dynamic Programming and the Calculus of Variations.* Academic Press, New York, NY, 1965.
5. Y. Gao, G. He, and C. Hou. On leveraging traffic predictability in active queue management. *Proc. IEEE INFOCOM*, 2002.
6. F. Kelly. Charging and rate control for elastic traffic. *European Trans. on Telecommunications*, 8:33–37, 1997.
7. F. Kelly. Mathematical modelling of the internet. *Mathematics Unlimited - 2001 and Beyond*, pages 685–702, 2001.
8. F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices,proportional fairness and stability. *J. of the Operational Research Society*, 49:237–252, 1998.
9. S. Kunniyur and R. Srikant. End-to-end congestion control: utility functions, random losses and ecn marks. *INFOCOM 2000*, 2000.
10. S. Kunniyur and R. Srikant. Analysis and design of an adaptive virtual queue algorithm for active queue management. *Proc. ACM Sigcomm*, 2001.
11. S. Kunniyur and R. Srikant. A time-scale decomposition approach to adaptive ecn marking. *IEEE Trans. on Automatic Control*, 2002.
12. S. Low. A duality model of tcp flow controls. *Proc. of ITC Sepcialist Seminar on IP Traffic Measurement, Modeling, and Management*, pages 18–20, 2000.
13. S. Low and D. Lapsley. Optimization flow control i: Basic algorithm and convergence. *IEEE Trans. on Networking*, 1999.
14. J. Padhey, *et al.* Modeling tcp throughput: A simple model and its empirical validation. *Proc. of SIFCOMM, Vancouver, Canada*, 1998.
15. F. Paganini. On the stability of optimization-based flow control. *Proc. of the American Control Conf.*, 2001.